March 1967

THE ASYMPTOTIC DISTRIBUTION OF A CLASS OF TWO-SAMPLE NON-LINEAR RANK ORDER STATISTICS IN THE NULL CASE[1]

Siegfried Schach

Technical Report No. 91

University of Minnesota
Minneapolis, Minnesota

[1]Research supported by the National Science Foundation under Grant No. GP-3816.

TABLE OF CONTENTS

I. Introduction

Let $X_1, X_2, \ldots, X_m$ and $Y_1, Y_2, \ldots, Y_n$ be independent samples from the same continuous distribution. Let $h_N(\cdot)$, $N = 1, 2, \ldots$ be a sequence of functions on $[-1,1]$, symmetric with respect to 0, and periodic with period 1, and let $R_1, R_2, \ldots, R_m$ be the ranks of the X-observations in the combined sample. In this report we find conditions under which a sequence of statistics of the form

$$T_N = \frac{1}{N} \sum_{i=1}^{m} \sum_{j=1}^{m} h_N\left(\frac{R_i - R_j}{N}\right) \tag{1}$$

converges in distribution as $N = m + n$ goes to infinity, and we derive the characteristic function of the limiting distribution.

II. An Example Studied by Wheeler and Watson.

Statistics of the form (1) arise naturally in non-parametric tests of equality of the underlying distributions, if two independent samples on the circle are given. In fact it can be shown that for a suitable group of transformations the locally most powerful invariant (non-parametric) test for detecting shift-alternatives of circular distributions is based on a statistic of this form wherever the distribution is sufficiently smooth.

Here we give one illustrative example: S. Wheeler and G. S. Watson (1964) proposed a non-parametric test for equality of two circular distributions which can be described as follows: Place the two samples on the circle, change the angles between "successive" observations in such a way that all are equal $(= \frac{2\pi}{N})$, compute the length R of the vector resultant of the "adjusted" X-observations and reject the Null hypothesis of equality if R is too large.

If points on the circle are considered as unit vectors on the complex plane, then, with the above terminology, the statistic R^2 is defined as follows:

$$R^2 = \sum_{j=1}^{m} e^{2\pi i R_j/N} \left(\overline{\sum_{k=1}^{m} e^{2\pi i R_k/N}} \right) = \sum_{j=1}^{m} \sum_{k=1}^{m} e^{2\pi i (R_j - R_k)/N}$$

$$= \sum_{j=1}^{m} \sum_{k=1}^{m} \cos 2\pi \left(\frac{R_j - R_k}{N} \right) \qquad \text{(since } R^2 \text{ is real)},$$

where the ranks R_j $(j = 1, 2, \ldots, m)$ are defined by choosing an arbitrary cut-off point and an arbitrary direction on the circle. Hence R^2/N is of the form (1) with $h_N(x) = \cos 2\pi x$, $N = 1, 2, \ldots$.

III. Limit Functions h and Equivalence Classes of Sequences $\{h_N\}$

Before any statement about the limiting behavior of a sequence $\{T_N, N = 1, 2, \ldots\}$ of the form (1) can be made, we have to make some assumptions about the sequence $\{h_N(\cdot), N = 1, 2, \ldots\}$. Actually $h_N(\cdot)$ has to be defined only at the points $\frac{K}{N}$, where $K = 0, \pm 1, \pm 2, \ldots, \pm N$, in order to make the expression (1) meaningful, but we assume that each $h_N(\cdot)$ is defined on $[-1,1]$ in order to facilitate the presentation of proofs and results. This is no real restriction on the class of statistics T_N. By $L_2[0,1]$ or L_2 we denote the space of square integrable functions on $[0,1]$. L_2 is a Hilbert space if the usual definition of inner product is used, and we denote by $||h||_{L_2}$, or simply $||h||$, the norm of an element of this Hilbert space.

Approximating sequences and equivalence classes of sequences. Throughout this paper we make the following two assumptions on a sequence written in the form $\{h_N(\cdot)\}$ or $\{h_N\}$:

(i) $h_N(\cdot)$ is defined on $[-1,1]$, symmetric with respect to 0 and periodic with period 1.

(ii) $h_N(\cdot)$ is a step function; it is constant on $(\frac{2K-1}{2N}, \frac{2K+1}{2N})$, $K = 0, \pm 1, \pm 2, \ldots$.[1]

[1]Note that the elements $h_N(\cdot)$ form a linear subspace of $L_2[-1,1]$.

Definition: We say $\{h_N\} \Rightarrow h$ ("$\{h_N\}$ converges to h" in a particular sense) if the following two conditions are satisfied:

(a) h is continuous, $\int_0^1 h(x)dx = 0$ and

$$||h - h_N||_{L_2} \to 0 \text{ as } N \to \infty.$$[2]

(b) $h_N(0) \to h(0)$ as $N \to \infty$.

We also assume that $\lambda_N = \frac{m_N}{N} \to \lambda$ as $N \to \infty$ $(0 \le \lambda \le 1)$ and that $\lim m_N \ge 3$.

On the basis of these assumptions we show that the limiting distribution of $\{T_N\}$, if it exists, is a function of $h(\cdot)$ only.

If we consider several sequences $\{h_N\}$, $\{g_N\}$, we sometimes denote the corresponding statistics by T_{h_N}, T_{g_N}, respectively.

Theorem 1:

(a) Let $\{h_N\}$ be a converging sequence $(h_N \Rightarrow h)$.

Then

$$ET_N = h(0)\,\lambda(1-\lambda) + o_N(1) \text{ and}$$

$$\text{var } T_N = 2\lambda^2 (1-\lambda)^2 ||h||^2 + o_N(1).$$

(b) If $\{g_N\}$, $\{h_N\}$ are two sequences satisfying (i) and (ii) and if $g_N(0) - h_N(0) \to 0$, $||g_N - h_N||_{L_2} \to 0$, i.e., if $\{g_N - h_N\} \Rightarrow 0$, then $E(T_{g_N} - T_{h_N})^2 \to 0$ as $N \to \infty$.

Proof of (a): We first give the proof under the assumption that for each N, $\sum_{i=1}^{N} h_N(\frac{i}{N}) = 0$.

For each N and each sample outcome we define a vector $z' = (z_1, \ldots, z_N)$ by

$$(2) \qquad z_i = \begin{cases} 1 & \text{if the } i^{th} \text{ value of the ordered combined sample comes from the first sample (X-sample).} \\ 0 & \text{if it comes from the second sample (Y-sample).} \end{cases}$$

[2]This obviously implies that $h(\cdot)$ is symmetric and periodic, since a subsequence of $h_N(\cdot)$ converges pointwise a.e. and since continuity implies uniqueness of the limit function.

With this notation T_N can obviously be written in the form

$$(3) \qquad T_N = \frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{N} h_N\left(\frac{i-j}{N}\right) z_i z_j .$$

We first compute $ET_N^2 = \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{i'=1}^{N} \sum_{j'=1}^{N} h_N\left(\frac{i-j}{N}\right) h_N\left(\frac{i'-j'}{N}\right) z_i z_j z_{i'} z_{j'}$.

$$\text{Obviously } Ez_i z_j z_{i'} z_{j'} = \begin{cases} \frac{m}{N} & \text{if all indices are equal,} \\ \frac{m}{N}\frac{m-1}{N-1} & \text{if any three indices are equal and the fourth is different or if any two pairs of indices are equal,} \\ \frac{m}{N}\frac{m-1}{N-1}\frac{m-2}{N-2} & \text{if there is exactly one pair of equal indices,} \\ \frac{m}{N}\frac{m-1}{N-1}\frac{m-2}{N-2}\frac{m-3}{N-3} & \text{if all four indices are different,} \end{cases}$$

for N large enough to make $m_N \geq 3$.

For convenience we use the symbol λ_{-k} to denote $\frac{m-k}{N-k}$, where we suppress the N on which λ_{-k} actually depends. Obviously $\lambda_{-k} \to \lambda$ as $N \to \infty$. To compute ET_N^2 we partition the set of quadruples of indices into seven subsets P_k:

$P_1 = \{(i,j,i',j') : i = j = i' = j'\}$,

$P_2 = \{(i,j,i',j') :$ three indices are equal, one is different$\}$,

$P_3 = \{(i,j,i',j') : i = j, \; i' = j', \; i \neq i'\}$,

$P_4 = \{(i,j,i',j') : i = i', \; j = j', \; i \neq j$ or $i = j', \; j = i', \; i \neq j\}$,

$P_5 = \{(i,j,i',j') : i = j$ or $i' = j'$, and all other indices are different$\}$,

$P_6 = \{(i,j,i',j') : i = i'$ or $i = j'$ or $j = j'$, and all other indices are different$\}$,

$P_7 = \{(i,j,i',j') :$ all indices are different$\}$.

On each of these partitions sets $Ez_i z_j z_{i'} z_{j'}$ is constant, hence it essentially suffices to compute

$$\sum_{P_k} h_N\left(\frac{i-j}{N}\right) h_N\left(\frac{i'-j'}{N}\right) \quad \text{for each } k.$$

For convenience we define $c_h^{(N)} = \frac{1}{N} \sum_{i=1}^{N} h_N(\frac{i}{N})^2$.

$$\sum_{P_1} h_N(\frac{i-j}{N})\, h_N(\frac{i'-j'}{N}) = N\, h_N(0)^2.$$

$$\sum_{P_2} h_N(\frac{i-j}{N})\, h_N(\frac{i'-j'}{N}) = 4 \sum_{i=1}^{N} \sum_{j\neq i}^{N} h_N(\frac{i-j}{N})\, h_N(0) = 4 \sum_{i=1}^{N} (-h_N(0)\, h_N(0))$$

$$= -4N\, h_N(0)^2.$$

(making use of the symmetry property and of the fact that $\sum_{1}^{N} h_N(\frac{i}{N}) = 0$)

$$\sum_{P_3} h_N(\frac{i-j}{N})\, h_N(\frac{i'-j'}{N}) = N(N-1)\, h_N(0)^2.$$

$$\sum_{P_4} h_N(\frac{i-j}{N})\, h_N(\frac{i'-j'}{N}) = 2 \sum_{i=1}^{N} \sum_{j\neq i} h_N(\frac{i-j}{N})^2 = 2 \sum_{i=1}^{N} (Nc_h^{(N)} - h_N(0)^2)$$

$$= 2N(Nc_h^{(N)} - h_N(0)^2).$$

$$\sum_{P_5} h_N(\frac{i-j}{N})\, h_N(\frac{i'-j'}{N}) = 2 \sum_{i=1}^{N} \sum_{j\neq i} \sum_{\substack{i'\neq i \\ i'\neq j}} h_N(\frac{i-j}{N})\, h_N(0) = -2(N-2)N\, h_N(0)^2.$$

$$\sum_{P_6} h_N(\frac{i-j}{N})\, h_N(\frac{i'-j'}{N}) = 4 \sum_{i=1}^{N} \sum_{j\neq i} \sum_{\substack{j'\neq j \\ j'\neq i}} h_N(\frac{i-j}{N})\, h_N(\frac{i'-j'}{N})$$

$$= 4 \sum_{i=1}^{N} \sum_{j\neq i} (-h_N(0) - h_N(\frac{i-j}{N}))h_N(\frac{i-j}{N})$$

$$= 4N\, h_N(0)^2 + 4N\, h_N(0)^2 - 4N^2\, c_h^{(N)}$$

$$= 8N\, h_N(0)^2 - 4N^2\, c_h^{(N)}.$$

$$\sum_{P_7} h_N(\frac{i-j}{N})\, h_N(\frac{i'-j'}{N}) = \sum_{i=1}^{N} \sum_{j\neq i} \sum_{\substack{i'\neq j \\ i'\neq i}} \sum_{\substack{j'\neq i' \\ j'\neq j \\ j'\neq i}} h_N(\frac{i-j}{N})\, h_N(\frac{i'-j'}{N})$$

$$= \sum_{i=1}^{N} \sum_{j\neq i} \sum_{\substack{i'\neq j \\ i'\neq i}} h_N(\frac{i-j}{N})(-h_N(0) - h_N(\frac{i'-j}{N}) - h_N(\frac{i'-i}{N}))$$

(continued)

$$= \sum_{i=1}^{N} \sum_{j\neq i} h_N(\tfrac{i-j}{N})(-(N-2)\, h_N(0) + 2h_N(0) + 2h_N(\tfrac{i-j}{N}))$$

$$= + (N-4)\, N\, h_N(0)^2 + 2N(Nc_h^{(N)} - h_N(0)^2)$$

$$= 2N^2\, c_h^{(N)} + N(N-6)\, h_N(0)^2.$$

Collecting terms and multiplying by $Ez_i z_j z_{i'} z_{j'}$ we get

$$N^2 ET_N^2 = N\, h_N(0)^2\lambda - 4N\, h_N(0)^2\lambda\lambda_{-1} + N(N-1)\, h_N(0)^2\lambda\lambda_{-1} + 2N(Nc_h^{(N)} - h_N(0)^2)\lambda\lambda_{-1}$$
$$-2(N-2)N\, h_N(0)^2\lambda\lambda_{-1}\lambda_{-2} + (8N\, h_N(0)^2 - 4N^2c_h^{(N)})\lambda\lambda_{-1}\lambda_{-2}$$
$$+ (2N^2c_h^{(N)} + N(N-6)\, h_N(0)^2)\lambda\lambda_{-1}\lambda_{-2}\lambda_{-3}.$$

Since $h_N(\cdot)$ is a step function which satisfies (i) and (ii) above we obtain the relation

$$c_h^{(N)} = \frac{1}{N}\sum_{i=1}^{N} h_N(\tfrac{i}{N})^2 = ||h_N||_{L_2}^2.$$

Upon dividing both sides of the above equation by N^2 and taking into account that $h_N(0) \Rightarrow h(0)$, we obtain

(4) $$ET_N^2 = h_N(0)^2(\lambda^2-2\lambda^3+\lambda^4) + ||h_N||^2(2\lambda^2-4\lambda^3+2\lambda^4) + o_N(1)$$

$$= [h(0)\lambda(1-\lambda)]^2 + 2[||h||_{L_2}\lambda(1-\lambda)]^2 + o_N(1).$$

Also

(5) $$ET_N = \frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N} h_N\,(\tfrac{i-j}{N})\, Ez_i z_j = \frac{1}{N}\sum_{i=1}^{N} h_N(0)\,\frac{m}{N} - \frac{N}{N}\, h_N(0)\,\frac{m}{N}\,\frac{m-1}{N-1}$$

$$= h_N(0)\,\frac{m}{N}\,(1 - \frac{m-1}{N-1}) = h(0)\lambda(1-\lambda) + o_N(1).$$

Hence

(6) $$\text{var } T_N = ET_N^2 - (ET_N)^2 = 2||h||_{L_2}^2\lambda^2(1-\lambda)^2 + o_N(1).$$

We now extend this result to the general case. Let $\bar{h}_N = \frac{1}{N} \sum_{i=1}^{N} h_N(\frac{i}{N})$, then the sequence $h'_N(\cdot) = h_N(\cdot) - \bar{h}_N$ satisfies the assumptions of the particular case treated so far.

But
$$|\bar{h}_N| = |\frac{1}{N} \sum_{i=1}^{N} h_N(\tfrac{i}{N}) - \int_0^1 h(x)dx| = |\int_0^1 (h_N(x) - h(x))dx|$$
$$\leq \int_0^1 |h_N(x) - h(x)|dx \to 0,$$

since convergence in L_2 implies convergence in L_1 on a finite measure space. Hence, applying (4) we get for the general case

$$(7) \qquad ET_N^2 = [(h_N(0) - \bar{h}_N)\,\lambda(1-\lambda)]^2 + 2[\lambda(1-\lambda)||h_N - \bar{h}_N||_{L_2}]^2 + o_N(1)$$
$$= [h_N(0)\,\lambda(1-\lambda)]^2 + o_N(1) + 2\lambda^2(1-\lambda)^2(||h_N||_{L_2}^2 - \bar{h}_N^2) + o_N(1)$$
$$= [h(0)\,\lambda(1-\lambda)]^2 + 2[\lambda(1-\lambda)||h||_{L_2}]^2 + o_N(1).$$

Similarly we have, making use of (5),

$$(8) \qquad ET_N = (h_N(0) - \bar{h}_N)\,\frac{m}{N}\,(1 - \frac{m-1}{N-1}) = h(0)\,\lambda(1-\lambda) + o_N(1),$$

and thus

$$(9) \qquad \mathrm{var}\; T_N = 2[\lambda(1-\lambda)||h||_{L_2}]^2 + o_N(1).$$

<u>Proof of (b)</u>: Replacing h_N by $g_N - h_N$ and T_N by $T_{g_N} - T_{h_N}$ we may apply the result derived under (a).

Hence from (7)

$$E(T_{g_N} - T_{h_N})^2 = [(g_N(0) - h_N(0))\,\lambda(1-\lambda)]^2 + 2[\lambda(1-\lambda)||g_N - h_N||_{L_2}]^2 + o_N(1)$$
$$= o_N(1).$$

This completes our proof.

If we define two sequences $\{g_N(\cdot)\}$ and $\{h_N(\cdot)\}$ satisfying $\{g_N - h_N\} \Rightarrow 0$ to be equivalent (the requirements of an equivalence relation are obviously satisfied), then equivalent sequences have identical limiting distributions, if a limiting distribution exists at all. We state this fact as a

Corollary:

If $\{h_N\}$ and $\{h_N'\}$ are sequences converging to the same h, then T_{h_N} and $T_{h_N'}$ either both converge in distribution to the same limit or neither of them converges in distribution.

Proof: $\{h_N\} \Rightarrow h$ and $\{h_N'\} \Rightarrow h$ implies $\{h_N - h_N'\} \Rightarrow 0$ and hence by Theorem 1 (b), $T_{h_N} - T_{h_N'} \xrightarrow{pr.} 0$ which implies the conclusion.

IV. (a) Asymptotic Distribution of T_N for h With Finite Fourier Expansion

Throughout this section we assume that $h(\cdot)$ has the Fourier expansion

$$(10) \qquad h(x) = \sum_{k=-K}^{K} d_k e^{2\pi ikx},$$

where K is arbitrary but finite. From our assumption $\int_0^1 h(x)dx = 0$ we get $d_0 = 0$.

Since $h(\cdot)$ is real we have $d_k = \bar{d}_{-k}$ for each k; but $h(\cdot)$ is also symmetric with respect to 0, hence

$$\sum_{k=-K}^{K} d_k e^{2\pi ikx} = h(x) = h(-x) = \sum_{k=-K}^{K} d_k e^{-2\pi ikx},$$

so that $d_k = d_{-k}$ (by uniqueness of expansion). It follows that d_k is real for all k. Combining the results we get

$$(11) \qquad d_0 = 0, \quad d_k = \bar{d}_k = d_{-k} \qquad k = \pm 1, \pm 2, \ldots, \pm K.$$

Matrix form of representation: By (3) T_N can be written in the form

$$(12) \qquad T_N = \frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{N} h_N\left(\frac{i-j}{N}\right) z_i z_j = \frac{1}{N} z'_N H_N z_N,$$

where $z_N = (z_1^{(N)}, z_2^{(N)}, \ldots, z_N^{(N)})'$ is the vector of indentically distributed, dependent random variables which are indicators of the X-sample. H_N is defined by

$$[H_N]_{r,s} = h_N\left(\frac{r-s}{N}\right) \qquad r = 1, 2, \ldots, N. \quad s = 1, 2, \ldots, N.$$

Diagonalization of H_N: H_N is a symmetric matrix (since $h_N(\cdot)$ is symmetric with respect to zero) which has the additional property that $[H_N]_{r,s}$ depends only on $(r - s) \bmod N$. Matrices with this latter property are called "circulant" matrices.

G. Wahba (1967) has shown that a unitary matrix W_N, which diagonalizes circulant matrices of order N, is given by the symmetric matrix W_N defined by

(13) $$[W_N]_{r,s} = \frac{1}{\sqrt{N}} e^{2\pi irs/N}.$$

By W_N^* we denote the adjoint of W_N, which is also equal to $\overline{W}_N$. Hence we get the relation

(14) $$\frac{1}{N} H_N = W_N D_N W_N^*,$$

where D_N is a diagonal matrix. Since H_N is symmetric, the elements of D_N are real.

If we set

(15) $$\eta_N = W_N^* z_N = (\eta_1^{(N)}, \eta_2^{(N)}, \dots, \eta_N^{(N)})'$$

we get

(16) $$T_N = \frac{1}{N} z_N' H_N z_N = z_N' W_N D_N W_N^* z_N = \bar{\eta}_N' D_N \eta_N = \sum_{\ell=1}^{N} d_\ell^{(N)} |\eta_\ell^{(N)}|^2$$

where $d_\ell^{(N)}$ are the diagonal elements of D_N $(\ell = 1, 2, \dots, N)$. Another way of writing this is

(17) $$T_N = \sum_{\ell=1}^{N} d_\ell^{(N)} [\mathrm{Re}(\eta_\ell^{(N)})^2 + \mathrm{Im}(\eta_\ell^{(N)})^2].$$

We now determine $d_\ell^{(N)}$:

Since $W_N^{-1} = W_N^*$ we get the relation (from (14))

(18) $$D_N = \frac{1}{N} W_N^* H_N W_N,$$

$$[H_N W_N]_{r,s} = \frac{1}{\sqrt{N}} \sum_{j=1}^{N} h_N\left(\frac{r-j}{N}\right) e^{2\pi ijs/N} = \frac{1}{\sqrt{N}} \sum_{j=1}^{N} h_N\left(\frac{j}{N}\right) e^{2\pi i(r-j)s/N},$$

by periodicity and symmetry of $h_N(\cdot)$.

Hence

$$\frac{1}{N}[W_N^* \; H_N \; W_N]_{t,s} = \frac{1}{N^2} \sum_{r=1}^{N} e^{-2\pi itr/N} e^{2\pi irs/N} \sum_{j=1}^{N} h_N(\tfrac{j}{N}) \; e^{2\pi ijs/N}$$

(by symmetry of h_N)

$$= \begin{cases} \frac{1}{N} \sum_{j=1}^{N} h_N(\frac{j}{N}) \; e^{2\pi is \frac{j}{N}} & \text{if } t = s, \\ 0 & \text{if } t \neq s, \end{cases}$$

using the orthogonality property of W_N.

It is easily recognized that

(19) $$d_\ell^{(N)} = \ell^{th} \text{ Fourier coefficient of } h_N(\cdot), \qquad \ell \leq N.$$

<u>Lemma 1</u>: If $\{h_N\} \Rightarrow h$, then $d_\ell^{(N)} \to d_\ell$ as $N \to \infty$, $\ell = 1, 2, \ldots$

<u>Proof</u>: If $(\cdot,\cdot)$ denotes the inner product in the Hilbert space $L_2[0,1]$, then

$$d_\ell = (h(x), e^{2\pi i\ell x}) \qquad \text{and} \qquad d_\ell^{(N)} = (h_N(x), e^{2\pi i\ell x}).$$

By the Schwarz inequality

$$|d_\ell - d_\ell^{(N)}|^2 = |(h(x) - h_N(x), e^{2\pi i\ell x})|^2 \leq ||h - h_N||^2_{L_2} ||e^{2\pi i\ell x}||^2 \to 0$$

as $N \to \infty$.

Equations (16) and (17) give a first clue as to what the limiting distribution of T_N might look like. The $\eta_\ell^{(N)}$'s are linear functions of the z_N's and can thus be expected to be asymptotically normal under quite general conditions. (We will come to this question later in this section.) The $d_\ell^{(N)}$ converge to known constants. This convergence, however, does not allow us to pass to the limit immediately, since it is not uniform in N. We shall see that in fact $d_\ell^{(N)} = d_{N-\ell}^{(N)}$. The next theorem will show how this obstacle can be overcome by exploiting more thoroughly the structure of W_N and by choosing a particularly suitable sequence $\{h_N\} \Rightarrow h$.

Theorem 2: Let $\{h_N\}$ be the sequence of step functions satisfying (i) and (ii), defined by $h_N(\frac{i}{N}) = h(\frac{i}{N})$, then $\{h_N\} \Rightarrow h$, and for $N > 2K$ we may write

$$T_{h_N} = 2\sum_{\ell=1}^{K} d_\ell |\eta_\ell^{(N)}|^2 = 2\bar{\eta}_N D_N' \eta_N \tag{20}$$

where $\eta_N = W_N^* z_N$, $\eta_\ell^{(N)} = \ell^{th}$ component of η_N, $[D_N']_{k,\ell} = \delta_{k\ell}' \cdot d_\ell$.

Proof: Since $h(\cdot)$ is uniformly continuous on $[0,1]$ it is obvious that $\{h_N\} \Rightarrow h$. By straightforward computation we obtain for the particular H_N with $[H_N]_{r,s} = h(\frac{r-s}{N})$:

$$[W_N^* H_N]_{r,s} = \frac{1}{\sqrt{N}} \sum_{t=1}^{N} e^{-2\pi irt/N} h(\frac{t-s}{N})$$

$$= \frac{1}{\sqrt{N}} \sum_{t=1}^{N} e^{-2\pi irt/N} \sum_{k=1}^{N} d_k (e^{2\pi ik(t-s)/N} + e^{-2\pi ik(t-s)/N})$$

$$[W_N^* H_N W_N]_{r,u} = \frac{1}{N} \sum_{s=1}^{N} e^{2\pi ius/N} \sum_{t=1}^{N} e^{-2\pi irt/N} \sum_{k=1}^{N} d_k (e^{2\pi ik(t-s)/N} + e^{-2\pi ik(t-s)/N}) \tag{21}$$

$$= \frac{1}{N} \sum_{k=1}^{N} d_k [\sum_{s=1}^{N} e^{2\pi i(u-k)s/N} \sum_{t=1}^{N} e^{-2\pi i(r-k)t/N} + \sum_{s=1}^{N} e^{2\pi i(u+k)s/N} \sum_{t=1}^{N} e^{-2\pi i(r+k)t/N}]$$

$$= \frac{1}{N} \sum_{k=1}^{N} d_k (\delta(u,r,k) + \delta'(u,r,k)), \text{ say.}$$

Now

$$\delta(u,r,k) = \begin{cases} N^2 & \text{if } k = u \text{ and } k = r, \\ 0 & \text{otherwise.} \end{cases}$$

$$\delta'(u,r,k) = \begin{cases} N^2 & \text{if } k = N-u \text{ and } k = N-r, \\ 0 & \text{otherwise.} \end{cases}$$

Hence

$$\frac{1}{N} W_N^* H_N W_N = \begin{pmatrix} d_1 & & & & & & & & & & \\ & d_2 & & & & & & & & & \\ & & \ddots & & & & & & & & \\ & & & d_K & & & & & & & \\ & & & & 0 & & & & & & \\ & & & & & \ddots & & & & & \\ & & & & & & 0 & & & & \\ & & & & & & & d_K & & & \\ & & & & & & & & \ddots & & \\ & & & & & & & & & d_2 & \\ & & & & & & & & & & d_1 \\ & & & & & & & & & & & 0 \end{pmatrix} = D_N,$$

where empty places are to indicate zeros.

From this we get the representation

$$T_N = \frac{1}{N} z_N' H_N z_N = z_N' W_N D_N W_N^* z_N = \sum_{\ell=1}^{K} d_\ell (|\eta_\ell^{(N)}|^2 + |\eta_{N-\ell}^{(N)}|^2),$$

where $\eta_\ell^{(N)} = \ell^{\text{th}}$ component of η_N, $\eta_N = W_N^* z_N$.

Now $|\eta_\ell^{(N)}|^2 + |\eta_{N-\ell}^{(N)}|^2 = (\mathrm{Re}\eta_\ell^{(N)})^2 + (\mathrm{Im}\eta_\ell^{(N)})^2 + (\mathrm{Re}\eta_{N-\ell}^{(N)})^2 + (\mathrm{Im}\eta_{N-\ell}^{(N)})^2$.

It is easy to see, that $\eta_\ell^{(N)}$ and $\eta_{N-\ell}^{(N)}$ are conjugate complex, and hence $|\eta_\ell^{(N)}|^2 + |\eta_{N-\ell}^{(N)}|^2 = 2|\eta_\ell^{(N)}|^2$. Hence we finally get

$$T_N = 2 \sum_{\ell=1}^{K} d_\ell |\eta_\ell^{(N)}|^2.$$

<u>Distribution of $\eta_\ell^{(N)}$</u>: From now on we assume that $N > 2K$, which is no real restriction, since we are interested in the limiting distribution of $\eta_\ell^{(N)}$ only.

<u>Theorem 3</u>:

(a) For each $\ell < N/2$ we have

(22) $$E\eta_\ell^{(N)} = 0.$$

(23) $$\mathrm{var}\ \mathrm{Re}(\eta_\ell^{(N)}) = \mathrm{var}\ \mathrm{Im}(\eta_\ell^{(N)}) = \frac{1}{2}\frac{m}{N}\left(1 - \frac{m-1}{N-1}\right) \to \frac{1}{2}\lambda(1-\lambda), \quad \text{as } N \to \infty.$$

(24) $$\mathrm{cov}\,(\mathrm{Re}(\eta_\ell^{(N)}), \mathrm{Im}(\eta_\ell^{(N)})) = 0.$$

(b) If $0 < \lambda < 1$, then $\mathrm{Re}(\eta_\ell^{(N)})$ and $\mathrm{Im}(\eta_\ell^{(N)})$ are asymptotically normally distributed with means and variances given by (22), (23).

<u>Proof</u>:

(a) $$E\eta_\ell^{(N)} = \sum_{r=1}^{N} \frac{1}{\sqrt{N}} e^{-2\pi i\ell r/N} E_{z_r} = \frac{m}{N}\frac{1}{\sqrt{N}} \sum_{r=1}^{N} e^{-2\pi i\ell r/N} = 0,$$

since the ℓ^{th} row of W_N^* is orthogonal to the N^{th}, which has the form $\frac{1}{\sqrt{N}}(1, 1, 1, \ldots, 1)$.

$$\begin{aligned}
\mathrm{var}\ \mathrm{Re}(\eta_\ell^{(N)}) &= E\ \mathrm{Re}(\eta_\ell^{(N)})^2 \\
&= \frac{1}{4N} E\Big[\sum_{r=1}^{N} e^{2\pi ir\ell/N} z_r + \sum_{s=1}^{N} e^{-2\pi is\ell/N} z_s\Big]^2 \\
&= \frac{1}{4N}\sum_{r=1}^{N}\sum_{s\neq r}(e^{2\pi ir\ell/N} + e^{-2\pi ir\ell/N})(e^{2\pi is\ell/N} + e^{-2\pi is\ell/N})\lambda\lambda_{-1} \\
&\quad + \frac{1}{4N}\sum_{r=1}^{N}(e^{2\pi ir\ell/N} + e^{-2\pi ir\ell/N})^2\lambda \\
&= \frac{1}{4N}\sum_{r=1}^{N}(e^{2\pi ir\ell/N} + e^{-2\pi ir\ell/N})^2\lambda(1-\lambda_{-1}) \\
&= \frac{\lambda(1-\lambda_{-1})}{4N}\begin{cases} 0 + 2N + 0 & \text{if } 2\ell \neq N, \\ N + 2N + N & \text{if } 2\ell = N. \end{cases}
\end{aligned}$$

$$\begin{aligned}
E|\eta_\ell^{(N)}|^2 &= \frac{1}{N} E\Big[\sum_{r=1}^{N}\sum_{s=1}^{N} e^{-2\pi i\ell r/N} e^{2\pi i\ell s/N} z_r z_s\Big] \\
&= \frac{1}{N}\sum_{r=1}^{N}\sum_{s\neq r} e^{-2\pi i\ell r/N} e^{2\pi i\ell s/N}\lambda\lambda_{-1} + \frac{1}{N}\sum_{r=1}^{N} e^{0}\lambda \\
&= \frac{1}{N}(-\lambda\lambda_{-1}N + \lambda N) = \lambda(1-\lambda_{-1}),
\end{aligned}$$

hence

$$\mathrm{var}\ \mathrm{Im}(\eta_\ell^{(N)}) = E|\eta_\ell^{(N)}|^2 - \mathrm{var}\ \mathrm{Re}(\eta_\ell^{(N)}) = \frac{1}{2}\lambda(1-\lambda_{-1}).$$

A similar straightforward computation shows that $\mathrm{Re}(\eta_\ell^{(N)})$ and $\mathrm{Im}(\eta_\ell^{(N)})$ are uncorrelated.

(b) We prove the result for $\alpha_N = \mathrm{Re}(\eta_\ell^{(N)})$, ℓ fixed. The proof for the imaginary part can be given in the same way.

Asymptotic normality follows easily from a very general result by Hájek (1961).

From his Theorems 4.1 and 4.2 we compile the following result:

Let $\{a_{Ni},\ i = 1, 2,\ldots,N\}$ and $\{b_{Ni},\ i = 1, 2,\ldots,N\}$ be double sequences of real numbers. Let $(R_{N1}, R_{N2},\ldots,R_{NN})$ be a random vector which assumes the $N!$ permutations of $(1, 2,\ldots,N)$ with equal probabilities. Set $\bar{a}_N = \frac{1}{N}\sum_{i=1}^{N} a_{Ni}$, $\bar{b}_N = \frac{1}{N}\sum_{i=1}^{N} b_{Ni}$.

Assume that

$$(25)\qquad \lim_{N\to\infty} \frac{\max\limits_{1\le i\le N} (a_{Ni} - \bar{a}_N)^2}{\sum\limits_{i=1}^{N} (a_{Ni} - \bar{a}_N)^2} = 0,$$

$$(26)\qquad \lim_{N\to\infty} \frac{\max\limits_{1\le i\le N} (b_{Ni} - \bar{b}_N)^2}{\sum\limits_{i=1}^{N} (b_{Ni} - \bar{b}_N)^2} = 0,$$

$$(27)^{1}\qquad \left[\lim_{N\to\infty} \frac{k_N}{N} = 0\right] \Rightarrow \lim_{N\to\infty} \frac{\max\limits_{1\le i_1 <\ldots< i_{k_N}\le N} \sum\limits_{\alpha=1}^{k_N} (a_{Ni_\alpha} - \bar{a}_N)^2}{\sum\limits_{i=1}^{N} (a_{Ni} - \bar{a}_N)^2} = 0.$$

Under these conditions $\alpha_N = \sum_{i=1}^{N} b_{Ni} a_{NR_i}$ is asymptotically normally distributed with mean $E\alpha_N$ and variance $\sigma^2(\alpha_N)$.

In our case we have

$$(28)\qquad \alpha_N = \frac{1}{\sqrt{N}} \sum_{j=1}^{N} \cos\,(2\pi j\ell/N) z_j .$$

---

[1] Here, of course, "$\Rightarrow$" stands for logical implication.

If we set

$$(29)\qquad a_{Nj} = \begin{cases} 1 & \text{if } 1 \le j \le m_N, \\ 0 & \text{if } m_N + 1 \le j \le N, \end{cases}$$

and

$$(30)\qquad b_{Nj} = \frac{1}{\sqrt{N}} \cos(2\pi j\ell/N),$$

Then it is easy to see that α_N defined by (28) has the form $\alpha_N = \sum_{j=1}^{N} b_{Nj} a_{NR_j}$.

Now we have to check conditions (25), (26), and (27).

$\max_j (a_{Nj} - \bar{a}_N)^2 \le 1$, $\sum_{j=1}^{N} (a_{Nj} - \bar{a}_N)^2 = m_N(1-\lambda_N)^2 + (N-m_N)\lambda_N^2 = N\lambda_N(1-\lambda_N)$ $\sim N\lambda(1-\lambda) \to \infty$, hence (25) is satisfied. $b^2_{Nj} \le \frac{1}{N}$ for all j, $(\bar{b}_N = 0)$ $\sum_{j=1}^{N} b^2_{Nj} = \frac{1}{N}\frac{1}{4}\sum_{j=1}^{N}(e^{2\pi ij\ell/N} + e^{-2\pi ij\ell/N})^2 = \frac{1}{4N}(0 + 2N + 0) = \frac{1}{2}$, this implies (26)

To check (27) let $\frac{k_N}{N} < \delta$ for $N \ge N_\delta$. Then $\sum_{\alpha=1}^{k_N}(a_{Nj_\alpha} - \bar{a}_N)^2 \le k_N < \delta N$ for $N \ge N_\delta$ and all indices $j_1 < j_2 < \ldots < j_{k_N}$. Since $\sum_{j=1}^{N}(a_{Nj} - \bar{a}_N)^2 \sim N\lambda(1-\lambda)$, we get $\dfrac{\sum_{\alpha=1}^{k_N}(a_{Nj_\alpha} - \bar{a}_N)^2}{\sum_{j=1}^{N}(a_{Nj} - \bar{a}_N)^2} < \dfrac{\delta}{\lambda(1-\lambda)}$ for N large enough. This shows that (27) is satisfied and completes our proof.

We are now going to extend the above result to any finite number of components $\eta_\ell^{(N)}$ $(\ell = 1, 2, \ldots, L)$, but before we state our theorem we prove a useful

<u>Lemma 2</u>: Let $w_\ell^{(N)} = \alpha_\ell^{(N)} + i\beta_\ell^{(N)}$ (α's and β's real) be the first L row vectors of W_N^*, (L fixed). Then any linear combination double sequence $\{b_{Nj}\}$ of the form $b_{Nj} = \sum_{\ell=1}^{L} c_\ell \alpha_{\ell j}^{(N)} + \sum_{\ell=1}^{L} d_\ell \beta_{\ell j}^{(N)}$ satisfies (26).

Proof: First note that for $N > 2L$ we have $\bar{\alpha}_\ell^{(N)} = \frac{1}{N} \sum_{j=1}^{N} \alpha_{\ell j}^{(N)} = 0$, $\bar{\beta}_\ell^{(N)} = \frac{1}{N} \sum_{j=1}^{N} \beta_{\ell j}^{(N)} = 0$, $\sum_{j=1}^{N} \alpha_{\ell j}^{(N)} \alpha_{\ell' j}^{(N)} = \sum_{j=1}^{N} \beta_{\ell j}^{(N)} \beta_{\ell' j}^{(N)} = 0$ for $1 \le \ell < \ell' \le L$

and $\sum_{j=1}^{N} \alpha_{\ell j}^{(N)} \beta_{\ell' j}^{(N)} = 0$ for $1 \le \ell \le \ell' \le L$ (i.e., all the real and complex components of W_N^* are orthogonal if $L < N/2$. We give the proof for a weighted sum of two double sequences $c_1\alpha_{1j}^{(N)} + c_2\alpha_{2j}^{(N)}$. The general case can be proved in the same way.

Consider $c_1\alpha_{1j}^{(N)} + c_2\alpha_{2j}^{(N)}$, $c_1^2 + c_2^2 > 0$.

$$\sum_{j=1}^{N} (c_1\alpha_{1j}^{(N)} + c_2\alpha_{2j}^{(N)})^2 = c_1^2 \sum_{j=1}^{N} \alpha_{1j}^{(N)2} + c_2^2 \sum_{j=1}^{N} \alpha_{2j}^{(N)2} = \frac{c_1^2}{2} + \frac{c_2^2}{2} > 0$$

$$\max_j (c_1\alpha_{1j}^{(N)} + c_2\alpha_{2j}^{(N)})^2 \le \max_j c_1^2\alpha_{1j}^{(N)2} + \max_j c_2^2\alpha_{2j}^{(N)2} + 2|c_1c_2|\max_j \alpha_{1j}^{(N)}\alpha_{2j}^{(N)}$$

$\le \frac{1}{N}(|c_1| + |c_2|)^2 \to 0$ as $N \to \infty$. Hence (26) follows.

Joint distribution of $\eta_\ell^{(N)}$, $\ell = 1, 2, \ldots, L$.

Theorem 4: Let $0 < \lambda < 1$. For any constant L the joint distribution of $(\mathrm{Re}\eta_1^{(N)}, \mathrm{Im}\eta_1^{(N)}, \mathrm{Re}\eta_2^{(N)}, \mathrm{Im}\eta_2^{(N)}, \ldots, \mathrm{Re}\eta_L^{(N)}, \mathrm{Im}\eta_L^{(N)})$ is asymptotically normal with mean vector 0 and covariance matrix $\Sigma = \frac{1}{2}\lambda(1-\lambda)I_{2L}$, where I_{2L} is the identity matrix of order $2L$.

Proof: According to a well-known theorem by H. Cramér it suffices to prove asymptotic normality for any linear combination $\zeta = \sum_{\ell=1}^{L} c_\ell \mathrm{Re}\eta_\ell^{(N)} + \sum_{\ell=1}^{L} c_\ell' \mathrm{Im}\eta_\ell^{(N)}$. But on account of Lemma 2 this follows directly from another application of Hájek's results. Computation of mean vector and covariance matrix is straightforward.

Asymptotic distribution of T_N.

Theorem 5: Let $0 < \lambda < 1$ and $\{h_N\} \Rightarrow h$. Under the assumptions of the present section, i.e., if $h(x) = \sum_{k=-K}^{K} d_k e^{2\pi ikx}$, the asymptotic distribution of

$T_N = \frac{1}{N}\sum_{i=1}^{m}\sum_{j=1}^{m} h_N\left(\frac{R_i - R_j}{N}\right)$ is the distribution of $\lambda(1-\lambda)\sum_{k=1}^{K} d_k u_k$, where u_k are independent χ^2 random variables with 2 d.f.

Proof: It is easy to see that if $F_N(x_1,\ldots,x_{2L}) \to F(x_1,\ldots,x_{2L})$ in distribution, where F is absolutely continuous, then the distribution of any quadratic form $x'Ax$ converges to the distribution of the quadratic form of the limit (the sets $x'Ax \le t$ are ellipses, and the F-measure of its boundaries are zero). Hence the result follows immediately from Theorem 2, (23), Theorem 4, and the corollary to Theorem 1.

Example: For the nonparametric test studied by Wheeler and Watson (1964) we have $h(x) = \cos 2\pi x$. Hence T_N tends in distribution to the exponential law with density $\theta e^{-\theta t}$, $t \ge 0$, and $1/\theta = \lambda(1-\lambda)$.

IV. (b) Extension of Results to a Class of Functions With Infinite Fourier Expansion

We now extend the results of the previous section to a class of functions h with infinitely many Fourier coefficients different from zero.

For a function $h(x) = \sum_{k=-K}^{K} d_k e^{2\pi ikx}$ we found the asymptotic distribution of T_N by first studying $2K$-dimensional linear functions of the z_N's, passing to the limit $(N \to \infty)$, and then deriving the distribution of the quadratic form.

In the present situation it seems to be natural that a similar path could be followed if we study "infinite-dimensional" linear functions of the z_N's taking values in some space and then try to pass to the limit.

A separable, infinite dimensional, Hilbert space will suit our purpose, as will be seen in the following theorems. In the appendix of this report we have presented a number of definitions and results on measures and convergence of measures on Hilbert spaces. A capital letter A refers to theorems and definitions stated in the appendix.

Under our assumptions on h (real, symmetric with respect to zero) we still have $d_k = d_{-k} = \bar{d}_k$ in the expansion

$$(31) \qquad h(x) = \sum_{-\infty}^{\infty} d_k e^{2\pi ikx}.$$

Throughout this section we make the following additional

<u>Assumption</u>:

$$(32) \qquad \sum_{-\infty}^{\infty} |d_k| = 2\sum_{1}^{\infty} |d_k| < \infty.$$

This assumption implies that $\sum_{-N}^{N} d_k \to h(0)$ as $N \to \infty$.

In this section it is convenient to use as approximating sequence the sequence $\{h_N\}$ satisfying (i) and (ii) and

$$(33) \qquad h_N(\tfrac{\ell}{N}) = \sum_{k=-N}^{N} d_k e^{2\pi ik\ell/N},$$

i.e., on the intervals of constancy h_N takes the values at $\frac{\ell}{N}$ of the N^{th} partial sum of the Fourier series.

<u>Lemma 3</u>: For the sequence $\{h_N\}$ satisfying (i) and (ii) defined by (33) we have $||h_N - h||_{L_2} \to 0$ and $h_N(0) \to h(0)$ as $N \to \infty$; i.e., $\{h_N\}$ is a legitimate approximation in the sense that $\{h_N\} \Rightarrow h$.

<u>Proof</u>: Since $\sum_{k=-N}^{N} d_k e^{2\pi ikx} \to h(x)$ uniformly in x ($h(\cdot)$ is continuous) we have $|h_N(x) - h^{(N)}(x)| < \epsilon$, all x, for $N \geq N_\epsilon$, where $h^{(N)}(x)$ is the approximating step function satisfying (i), (ii) and $h^{(N)}(\frac{i}{N}) = h(\frac{i}{N})$. Hence $||h_N - h^{(N)}||_{L_2} \to 0$ as $N \to \infty$. By uniform continuity of h we have $||h^{(N)} - h||_{L_2} \to 0$, and hence $||h_N - h||_{L_2} \to 0$ as $N \to \infty$.

We have already remarked that $h_N(0) \to h(0)$ because of assumption (32).

We now define a sequence of measures on the real Hilbert space H of real sequences with finite sum of squares.

As always in probability theory we assume that there exists an underlying probability space $(\Omega, \mathfrak{A}, P)$ where the sequences of i.i.d. random variables $X_1, X_2, \ldots$ ad inf. and $Y_1, Y_2, \ldots$ ad inf. (with continuous distribution function) are defined. We also assume that to each N there is defined an integer $m_N \le N$, such that $\lambda_N = \frac{m_N}{N} \to \lambda$ as $N \to \infty$, and $0 < \lambda < 1$.

Since we are only interested in certain rank order statistics, all our random variables are functions of the vectors z_N defined by (2). But it should be kept in mind that they are actually measurable functions on $(\Omega, \mathfrak{A}, P)$.

Any measurable transformation from the space of the z_N's to H induces a probability measure on H. We now define a sequence of transformations S_N of the z_N's and hence a sequence of measures on H:

Let V_N be the $N \times 2N$ matrix defined by

$$(34) \qquad [V_N]_{r,s} = \begin{cases} \mathrm{Re}\ [W_N]_{r,\frac{s+1}{2}} & \text{if } s \text{ is odd,} \\ \mathrm{Im}\ [W_N]_{r,\frac{s}{2}} & \text{if } s \text{ is even.} \end{cases}$$

$$\text{Set } \quad c_k^{(2N)} = \begin{cases} |d_\ell| + |d_{N-\ell}| & \text{where } \ell = [\frac{k+1}{2}], \quad 1 \le k \le 2[\frac{N}{2}], \\ 0 & \text{for } 2[\frac{N}{2}] < k \le 2N-2, \\ |d_N| & \text{for } k = 2N-1 \text{ and for } k = 2N. \end{cases}$$

Define $\bar{C}_{2N}$ to be the diagonal matrix with elements $c_k^{(2N)}$ $(k = 1, 2, \ldots, 2N)$. E.g., for N even we have

$$\bar{C}_{2N} = \begin{pmatrix} |d_1|+|d_{N-1}| & & & & & & & & & \\ & |d_1|+|d_{N-1}| & & & & & & & & \\ & & |d_2|+|d_{N-2}| & & & & & & & \\ & & & \ddots & & & & & & \\ & & & & |d_{N/2}|+|d_{N/2}| & & & & & \\ & & & & & 0 & & & & \\ & & & & & & \ddots & & & \\ & & & & & & & 0 & & \\ & & & & & & & & |d_N| & \\ & & & & & & & & & |d_N| \end{pmatrix}$$

If $x = (x_1, x_2, \ldots)$ is a generic element of the Hilbert space H, we define the sequence $\{S_N\}$ of mappings from Ω to H by the relation

$$(36) \qquad \omega \to z_N(\omega) \to x(\omega), \qquad \text{with} \begin{cases} (x_1, x_2, \ldots, x_{2N})' = \overline{C}_N V_N' z_N \\ x_i = 0 \quad \text{for} \quad i > 2N. \end{cases}$$

Let μ_N be the probability measure on H induced by S_N.

<u>Theorem 6</u>: The sequence $\{\mu_N\}$ of probability measures is compact.

<u>Proof</u>: We use Theorem A.2 for the proof.

Condition 1: This is obviously satisfied.

Condition 2: As a basis for H we take the vectors $(1,0,0,\ldots)$, $(0,1,0,0,\ldots), \ldots$. Now let $\varepsilon > 0$ be given. Take M_ε, even, such that $\sum_{k=\frac{1}{2}M_\varepsilon}^{\infty} |d_k| < \frac{\varepsilon}{4}$, which is possible by (32), and let $M > M_\varepsilon$ be arbitrary. Then for $N < M$ we have

$$\int_H r_M^2(x)\, d\mu_N(x) \le \begin{cases} 0 & \text{if } 2N < M, \\ 2|d_N|\frac{1}{2}\lambda(1-\lambda_{-1}) < \varepsilon & \text{for } M \le 2N < 2M, \text{ by (23).} \end{cases}$$

For $N \ge M$ we set $\xi_{2N} = V_N' z_N$ (this is the vector of real and imaginary parts of η_N) and use the fact that all the components $\xi_\ell^{(2N)}$ have the same expected square $\frac{1}{2}\lambda(1-\lambda_{-1})$ (by (23)). Hence

$$\begin{aligned} \int_H r_M^2(x)\, d\mu_N(x) &= \sum_{k=M}^{2N} (c_k^{(2N)})^2\, E(\xi_k^{(2N)})^2 \\ &= \frac{1}{2}\lambda(1-\lambda_{-1}) \sum_{k=M}^{2N} (c_k^{(2N)})^2 \\ &\le \frac{1}{2}\lambda(1-\lambda_{-1})[2 \sum_{\ell=[\frac{M+1}{2}]}^{[\frac{N+1}{2}]} (|d_\ell| + |d_{N-\ell}|) + 2|d_N|] \\ &\le \lambda(1-\lambda)[\sum_{\ell=[\frac{M}{2}]}^{\infty} |d_\ell| + \sum_{\ell=[\frac{M}{2}]}^{\infty} |d_\ell| + |d_N|] < \varepsilon. \end{aligned}$$

Hence for $M > M_\varepsilon$ we have $\sup_N \int_H r_M^2(x)\, d\mu_N(x) \le \varepsilon$, which completes our proof.

Lemma 4: Let $\{\mu_N\}$ be a compact sequence, let M be the closure (in the topology of weak convergence) of $\{\mu_N, N = 1, 2, \ldots\})$. Then the set of characteristic functionals $\{\chi(f,\mu): \mu \in M\}$ is a uniformly equicontinuous set of functions of f (on H). (For the definition of $\chi(f,\mu)$ see Appendix (A.3).)

Proof: Let $\epsilon > 0$ be given. By Theorem A.1 there exists a compact K_ϵ such that $\mu_N(K_\epsilon) \geq 1 - \frac{\epsilon}{4}$, all N. By Lemma A.1 any $\mu \in M$ also satisfies this relation, since it is the limit of a suitably chosen subsequence μ_{N_i}. Now K_ϵ is bounded by some constant $K > 0$. Let $||f - g|| < \frac{\epsilon}{2K}$. Then for any $\mu \in M$

$$|\chi(f,\mu) - \chi(g,\mu)| = \left|\int_H (e^{i(f,x)} - e^{i(g,x)}) d\mu(x)\right| \leq \int_{K_\epsilon} |e^{i(f,x)}| \, |1-e^{i(g-f,x)}| d\mu(x)$$

$$+ 2\int_{H-K_\epsilon} d\mu(x) \leq \int_{K_\epsilon} |(g-f,x)| d\mu(x) + \frac{\epsilon}{2} \leq \int_{K_\epsilon} ||f-g|| \; ||x|| \; d\mu(x) + \frac{\epsilon}{2}$$

$$\leq K \frac{\epsilon}{2K} + \frac{\epsilon}{2} = \epsilon.$$

Theorem 7: $\mu_N \to$ some probability measure μ as $N \to \infty$.

Proof: If $\mu_N \to \mu$ in the sense of weak convergence, then $\mu_N(R) = 1 = \int 1 d\mu_N \to \int 1 d\mu = \mu(R)$, thus μ has to be a probability measure. We show weak convergence by using Theorem A3.

Let $f^{(k)} = (f_1, f_2, \ldots, f_k, 0, 0, \ldots)$; then by a trivial extension of Theorem 4 we get

$$(37) \qquad \chi(f^{(k)}, \mu_N) \to e^{-\frac{1}{2}\frac{\lambda(1-\lambda)}{2}\sum_{i=1}^{k}|d_{[\frac{i+1}{2}]}|f_i^2} \quad \text{as } N \to \infty.$$

Let μ be any limit measure of a suitably chosen subsequence. Then by the definition of weak convergence

$$(38) \qquad \chi(f^{(k)}, \mu) = e^{-\frac{1}{2}\frac{\lambda(1-\lambda)}{2}\sum_{i=1}^{k}|d_{[\frac{i+1}{2}]}|f_i^2}.$$

Since the $f^{(k)}$'s are dense in H, and since the left hand side of (38) is continuous in its first argument we must have:

$$(39) \qquad \chi(f, \mu) = e^{-\frac{1}{2}\frac{\lambda(1-\lambda)}{2}\sum_{i=1}^{\infty}|d_{[\frac{i+1}{2}]}|f_i^2} \quad \text{all } f \in H.$$

(I.e., any limiting measure is normal with S-operator of the form

$$[S]_{r,s} = \frac{\lambda(1-\lambda)}{2} \delta_{r,s} |d_{[\frac{r+1}{2}]}|.)$$

Now let $f \in H$ and $\epsilon > 0$ be given. By the preceding Lemma 4 we know that $|\chi(f,\bar{\mu}) - \chi(f^{(k)},\bar{\mu})| < \frac{\epsilon}{3}$ uniformly in $\bar{\mu} \in M$, if only $||f - f^{(k)}|| \leq \delta_\epsilon$ for suitably chosen δ_ϵ. Fix such an $f^{(k)}$ and let N_ϵ be such that $|\chi(f^{(k)},\mu_N) - \chi(f^{(k)},\mu)| < \frac{\epsilon}{3}$ for $N \geq N_\epsilon$, and for the particular μ appearing in (38) and (39). Then for $N \geq N_\epsilon$ we get

$$\begin{aligned}
&|\chi(f,\mu_N) - \chi(f,\mu)| \\
&\quad \leq |\chi(f,\mu_N) - \chi(f^{(k)},\mu_N)| + |\chi(f^{(k)},\mu_N) - \chi(f^{(k)},\mu)| \\
&\qquad\qquad + |\chi(f^{(k)},\mu) - \chi(f,\mu)| \\
&\quad \leq \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon .
\end{aligned}$$

By Theorem A.3 we get the desired result.

<u>Theorem 8</u>: Under the assumptions of this section T_N converges in distribution to the distribution with characteristic function

$$(40) \qquad \varphi_T(t) = \prod_{k=1}^{\infty} (1 - 2i\lambda(1-\lambda)\, d_k t)^{-1}.$$

<u>Proof</u>:

Let $\delta_k := \begin{cases} +1, & \text{if } d_k \geq 0, \\ -1, & \text{if } d_k < 0. \end{cases}$

If $\eta_N = W_N^* z_N$ as previously, then for the components $\eta_\ell^{(N)}$ we have $\eta_{N-\ell}^{(N)} = \bar{\eta}_\ell^{(N)}$. With this notation a straightforward computation shows that

$$(41) \qquad \begin{aligned} T_N &= \frac{1}{N} z_N' H_N z_N = z_N' W_N D_N W_N^* z_N = \bar{\eta}_N' D_N \eta_N = \sum_{k=1}^{N} d_k^{(N)} |(\eta_N)|^2 \\ &= \sum_{k=1}^{N} d_k^{(N)} [\mathrm{Re}\,\eta_k^{(N)})^2 + (\mathrm{Im}\,\eta_k^{(N)})^2] = \sum_{k=1}^{N} \delta_k (c_{2k}^{(2N)})^2 [(\xi_{2k-1}^{(2N)})^2 + (\xi_{2k}^{(2N)})^2] \end{aligned}$$

where $d_k^{(N)}$ are the characteristic values of H_N obtained by diagonalizing H_N by the matrix W_N.

Combining (36) and (41) it is easy to see that $T_N(\omega)$ is a function of $S_N(\omega)$ and that on H T_N has the simple structure

$$(42) \qquad T_N = \sum_{k=1}^{N} \delta_{[\frac{k+1}{2}]} x_k^2 + \delta_N(x_{2N-1}^2 + x_{2N}^2) = \sum_{k=1}^{\infty} \delta_{[\frac{k+1}{2}]} x_k^2 \text{ a.e. } (\mu_N),$$

since $x_{N+1} = \ldots = x_{2N-2} = x_{2N+1} = \ldots = 0$ a.e. μ_N.

Now we know from weak convergence that

$$(43) \qquad \varphi_{T_N}(t) = E_{\mu_N} e^{it \sum_{k=1}^{N} \delta_{[\frac{k+1}{2}]} x_k^2} = E_{\mu_N} e^{it \sum_{k=1}^{\infty} \delta_{[\frac{k+1}{2}]} x_k^2}$$

$$\to E_{\mu} e^{it \sum_{1}^{\infty} \delta_{[\frac{k+1}{2}]} x_k^2} = \varphi_T(t) \quad \text{as} \quad N \to \infty.$$

This limit is continuous and hence T_N converges in distribution. We now evaluate $E_{\mu} e^{it \sum_{k=1}^{\infty} \delta_{[\frac{k+1}{2}]} x_k^2}$. From Theorem 7 we conclude that for each K

$$(44) \qquad E_{\mu_N} e^{it \sum_{k=1}^{2K} \delta_{[\frac{k+1}{2}]} x_k^2} \to \prod_{k=1}^{K} (1 - 2i\lambda(1-\lambda)d_k t)^{-1} = E_{\mu} e^{it \sum_{k=1}^{2K} \delta_{[\frac{k+1}{2}]} x_k^2}.$$

By the dominated convergence theorem we can pass to the limit in K and get

$$\varphi_T(t) = E_{\mu} e^{it \sum_{k=1}^{\infty} \delta_{[\frac{k+1}{2}]} x_k^2} = \prod_{k=1}^{\infty} (1 - 2i\lambda(1-\lambda)\, d_k t)^{-1}.$$

<u>Theorem 9</u>: Let h have a continuous derivative. Then T_N converges in distribution to a probability measure with characteristic function

$$\varphi(t) = \prod_{k=1}^{\infty} (1 - 2i\lambda(1-\lambda)\, d_k t)^{-1},$$

where $d_k = \int_0^1 e^{2\pi ikx} h(x)\, dx, \qquad k = 1, 2, \ldots$

<u>Proof</u>: It is a well-known fact in Fourier-analysis that under the above conditions $|d_k| \leq \frac{M}{k^2}$ for some constant M. Hence the sequence of Fourier coefficients converges absolutely and the result follows from Theorem 8.

Appendix

Definitions and Results about Measures and Weak Convergence of Measures on Separable Hilbert Spaces

In this section we present a few definitions and results on convergence of measures on Hilbert spaces. Most of the results follow directly from the definitions or can be found in Y. V. Prokhorov (1956).

Throughout this appendix we assume that the Hilbert space H is real and separable.

Let $\mathcal{B}$ be the σ-field of subsets of H generated by the class of continuous linear functionals on H (i.e., the minimal σ-field with respect to which all elements of the dual space are measurable), then $(H, \mathcal{B})$ is a measurable space. Any countably additive nonnegative set function $m(\cdot)$ defined on $\mathcal{B}$ is called a measure on H.

We always assume that $m(H) < \infty$.

<u>Definition</u>: A sequence of measures $\{\mu_N\}$ converges weakly to a measure μ (in symbols: $\mu_N \xrightarrow{\text{weakly}} \mu$), if

(A1) $$\int_H f(x)\, d\mu_N(x) \to \int_H f(x)\, d\mu(x)$$

for all bounded continuous functions f on H.

This type of convergence is usually called "convergence in distribution" if H is finite-dimensional.

Convergence in distribution on finite dimensional spaces can be metrized. The so-called Lévy-metric has the property that convergence of measures in this metric is equivalent to convergence in distribution.

A similar metric has been defined by Prokhorov for the class of finite measures on $(H, \mathcal{B})$. Convergence in this metric is the same as weak convergence. With this metric the class $M(H)$ of finite measures on $(H, \mathcal{B})$ becomes a complete separable metric space.

Lemma A.1: $\mu_N \xrightarrow{\text{weakly}} \mu$ implies $\overline{\lim_N}\, \mu_N(F) \leq \mu(F)$ for any closed set $F \subset H$.

Of great interest is the characterization of the compact subsets of $M(H)$, i.e., sets which can be covered by a finite ϵ-net for every $\epsilon > 0$.

Theorem A.1: A set $M' \subset M(H)$ is compact if and only if

(1) $\sup_{\mu \in M'} \mu(H) < \infty$ and

(2) for any $\epsilon > 0$ there exists a compact $K_\epsilon \subset H$ such that $\mu(H - K_\epsilon) \leq \epsilon$ for every $\mu \in M'$.

A useful sufficient condition for compact subsets M' is given by the following theorem. Since H is separable there exists a countable complete orthonormal set $\{e_i\}$ of vectors. We assume that such a system has been chosen and define

$$r_N^2(x) = \sum_{i=N}^{\infty} (x, e_i)^2. \tag{A2}$$

Then we can state this

Theorem A.2: A set of measures $M' \subset M(H)$ for which

(1) $\sup_{\mu \in M'} \mu(H) < \infty$, and

(2) $\lim_{N \to \infty} \sup_{\mu \in M'} \int_H r_N^2(x)\, d\mu(x) = 0$,

is compact.

A powerful tool for analyzing convergence in distribution of measures on finite dimensional spaces is the characteristic function. A corresponding transformation can be defined on Hilbert spaces.

Definition: The function $\chi(\cdot, \mu)$ defined for any bounded linear functional f on H by the equation

$$\chi(f, \mu) = \int_H e^{i(f,x)} d\mu(x) \tag{A3}$$

is called characteristic functional of the measure μ. Here $(\cdot,\cdot)$ is the inner product on H.

Since μ is finite it is obvious that the integral exists. The name characteristic functional is justified by the fact that a measure is uniquely determined by its characteristic functional.

A characteristic functional is continuous in f.

By the definition of weak convergence we have, of course,

$$\mu_N \xrightarrow{\text{weakly}} \mu \Rightarrow \chi(f,\mu_N) \to \chi(f,\mu), \qquad \text{every } f \in H.$$

The converse $[\chi(f,\mu_N) \to \chi(f)$, continuous$] \Rightarrow \mu_N \xrightarrow{\text{weakly}} \mu$ is unfortunately not true for infinite-dimensional Hilbert spaces (and the usual topology on H). We have however the weaker

Theorem A.3: If $\{\mu_N, \ N = 1, 2, \dots\}$ is compact and $\chi(f,\mu_N) \to \chi(f)$ for every $f \in H$, then for some $\mu \in M(H)$

$$\mu_N \xrightarrow{\text{weakly}} \mu \quad \text{and} \quad \chi(f) = \chi(f,\mu).$$

References

Hájek, J. (1961). "Some extensions of the Wald - Wolfowitz - Noether theorem." Ann. Math. Stat. Vol. 32, pp. 506-23.

Prokhorov, Y. V. (1956). "Convergence of random processes and limit theorems in probability theory." Theory of Probability and its Applications, Vol. 1, pp. 158-214.

Wahba, G. (1967). "On the joint distribution of circulant forms in jointly stationary time series." Stanford University, Department of Statistics, Technical Report No. 17.

Wheeler, S. and Watson, G. S. (1964). "A distribution-free two-sample test on a circle." Biometrika, Vol. 51, pp. 256-58.